

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-306599

(43)Date of publication of application : 02.11.2001

(51)Int.Cl.

G06F 17/30

G06F 12/00

H04N 5/76

H04N 5/765

H04N 5/781

(21)Application number : 2000-120657

(71)Applicant : NIPPON TELEGR & TELEPH CORP
<NTT>

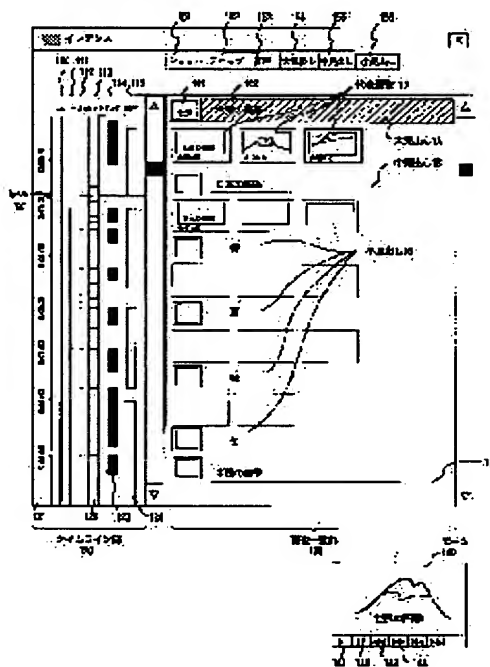
(22)Date of filing : 21.04.2000

(72)Inventor : TANIGUCHI YUKINOBU
KOJIMA HARUHIKO(54) METHOD AND DEVICE FOR HIERARCHICALLY MANAGING VIDEO, AND RECORDING
MEDIUM RECORDED WITH HIERARCHICAL MANAGEMENT PROGRAM

(57)Abstract:

PROBLEM TO BE SOLVED: To describe the logical hierarchical structure of video and to apply the logical structure with the intuitive operation of an operator.

SOLUTION: An index item list is prepared by enumerating and arraying plural logical units at random while defining a tag, display starting time and an index character corresponding to the logical unit of video as one item and the aligned index item list is displayed, by aligning the index items in the order of small index levels or starting time extracted corresponding to the tag.



LEGAL STATUS

[Date of request for examination] 25.12.2001

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3574606

[Date of registration] 09.07.2004

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's]

decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2001-306599
(P2001-306599A)

(43) 公開日 平成13年11月2日 (2001.11.2)

(51) Int.Cl. ⁷	識別記号	F I	テーマコード (参考)	
G 0 6 F 17/30	3 1 0	G 0 6 F 17/30	3 1 0 B	5 B 0 7 5
	1 7 0		1 7 0 D	5 B 0 8 2
	2 1 0		2 1 0 C	5 C 0 5 2
	4 1 4		4 1 4 A	
	4 1 9		4 1 9 A	

審査請求 未請求 請求項の数12 OL (全 17 頁) 最終頁に続く

(21) 出願番号 特願2000-120657(P2000-120657)

(22) 出願日 平成12年4月21日 (2000.4.21)

(71) 出願人 000004226

日本電信電話株式会社

東京都千代田区大手町二丁目3番1号

(72) 発明者 谷口 行信

東京都千代田区大手町二丁目3番1号 日
本電信電話株式会社内

(72) 発明者 児島 治彦

東京都千代田区大手町二丁目3番1号 日
本電信電話株式会社内

(74) 代理人 100087848

弁理士 小笠原 吉義 (外1名)

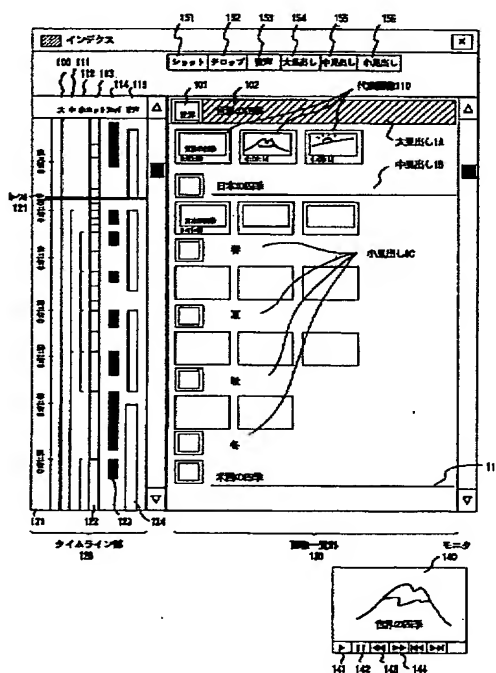
最終頁に続く

(54) 【発明の名称】 映像の階層的管理方法および階層的管理装置並びに階層的管理プログラムを記録した記録媒体

(57) 【要約】

【課題】 本発明は、映像の論理的な階層構造を記述可能にし、オペレータの直感的な操作によって論理構造を付与可能にすることを目的としている。

【解決手段】 映像の論理単位に対応する、タグと表示開始時刻と、見出し文字とを1つの項目として複数の論理単位を列記して順不同に配列された見出し項目リストを作成し、上記タグに対応して抽出される見出しレベルの小さい順にかつ開始時刻順に見出し項目を整列させて、整列された見出し項目リストを表示する。



【特許請求の範囲】

【請求項1】 映像を構成する個々の論理単位の内容を、コンピュータを用いて、表現し管理する映像管理方法において、

当該個々の論理単位についての、当該論理単位に対応するタグと当該論理単位に対応する映像を表示する開始時刻と当該論理単位に対応して存在し得る見出し文字列とを、1つの見出し項目として指定する処理を繰り返し実行して、見出し項目が配列された見出し項目リストを作成する過程と、

当該見出し項目リスト内の各見出し項目について、各見出し項目の前記タグにもとづいて、当該見出し項目に割りつけられる見出しレベルを取得する過程と、

前記見出し項目リスト内の各見出し項目について、各見出し項目を、前記開始時刻の小さい順にかつ開始時刻が同じであれば見出しレベルの小さい順に整列させる過程と、

当該整列された見出し項目について、当該整列された順に夫々の見出し項目についての見出しレベルの大小関係を表示する過程とを有することを特徴とする映像の階層的管理方法。

【請求項2】 請求項1記載の映像の階層的管理方法であって、

見出し項目リストに記録された開始時刻と見出しレベルを参照して、見出し項目をノードとする階層構造に変換し、各見出し項目を階層的に表示すること、
ことを特徴とする映像の階層的管理方法。

【請求項3】 請求項1または2記載の映像の階層的管理方法であって、

映像の論理構造の終了時刻を更に見出し項目に指定することを特徴とする映像の階層的管理方法。

【請求項4】 請求項1または2または3記載の映像の階層的管理方法であって、

映像から抜き出した複数の代表画像を見出し文字列と同時に一覧表示することを特徴とする映像の階層的管理方法。

【請求項5】 請求項4記載の映像の階層的管理方法であって、

時間軸上に代表画像と一対一に対応付けられた帯をタイムライン表示し、

タイムライン表示された帯を選択すると、対応する代表画像が選択されて当該代表画像が選択されたことが可視的に表示されることを特徴とする映像の階層的管理方法。

【請求項6】 請求項4または5記載の映像の階層的管理方法であって、

前記見出し項目に対応する代表画像と関連付けて、当該見出し項目に関係する情報を管理することを特徴とする映像の階層的管理方法。

【請求項7】 請求項1記載の映像の階層的管理方法で

あって、

前記見出し項目が整列された整列済み見出し項目リストをインデクスデータとして、映像データと関連付けて蓄積しておく、

複数の当該整列済み見出し項目リストの少なくとも一つを選択させ、当該選択された整列済み見出し項目リスト内の見出し項目に対応する論理単位について、当該論理単位の開始時間から蓄積された映像データを再生することを特徴とする映像の階層的管理方法。

【請求項8】 請求項1記載の映像の階層的管理方法であって、

前記整列済み見出し項目リストとともに、検索用インデクスを蓄積しておく、

利用者が検索キー情報を指示したことを契機として、前記検索キー情報と一致あるいは類似する情報を含むインデクスを列挙するようにし、

次いで、当該インデクスに対応する見出し項目と、表示のための時期が時間的に重なる見出し項目を、前記複数の整列済み見出し項目リストの中から検索し、

検索された当該見出し項目の群を、当該見出し項目を含む前記整列済み見出し項目リストにおける目次情報と一緒に表示することを特徴とする映像の階層的管理方法。

【請求項9】 映像を構成する個々の論理単位の内容を、コンピュータを用いて、表現し管理する映像管理装置において、

当該個々の論理単位についての、当該論理単位に対応するタグと当該論理単位に対応する映像を表示する開始時刻と当該論理単位に対応して存在し得る見出し文字列とを、1つの見出し項目として指定して、複数の当該見出し項目が順不同に配列された見出し項目リストと、
前記論理単位に対応するタグに対応して割りつけられる見出しレベルを記述した対応表と、

をそなえると共に、
前記見出し項目リスト内の各見出し項目について、各見出し項目の前記タグにもとづいて、前記対応表を索引して、当該見出し項目に割りつけられる見出しレベルを取得する処理手段と、

前記見出し項目リスト内の各見出し項目について、各見出し項目を、前記開始時刻の小さい順にかつ開始時刻が同じであれば見出しレベルの小さい順に整列させる処理手段と、

当該整列された見出し項目について、当該整列された順に夫々の見出し項目についての見出しレベルの大小関係を表示する処理手段とを有することを特徴とする映像の階層的管理装置。

【請求項10】 請求項9記載の映像の階層的管理装置であって、
前記見出し項目が整列された整列済み見出し項目リストをインデクスデータとして、映像データと関連付けて蓄積した蓄積手段と、

複数の当該整列済み見出し項目リストの少なくとも一つを選択させた際に、当該選択された整列済み見出し項目リスト内の見出し項目に対応する論理単位について、当該論理単位の開始時間から蓄積された映像データを再生する手段とを有することを特徴とする映像の階層的管理装置。

【請求項11】 映像を構成する個々の論理単位の内容を、コンピュータを用いて、表現し管理する映像管理方法について、当該映像管理方法を記述したプログラムを記録した記録媒体において、

前記映像管理方法を記述したプログラムが前記個々の論理単位についての、当該論理単位に対応するタグと当該論理単位に対応する映像を表示する開始時刻と当該論理単位に対応して存在し得る見出し文字列とを、1つの見出し項目として指定する処理を繰り返し実行して得られた、見出し項目が順不同に配列された見出し項目リストを取込み、

当該見出し項目リスト内の各見出し項目について、各見出し項目の前記タグにもとづいて、当該見出し項目に割りつけられる見出しレベルを取得する処理過程と、

前記見出し項目リスト内の各見出し項目について、各見出し項目を、前記開始時刻の小さい順にかつ開始時刻が同じであれば見出しレベルの小さい順に整列させる処理過程と、

当該整列された見出し項目について、当該整列された順に夫々の見出し項目についての見出しレベルの大小関係を表示する処理過程とを順に実行するよう記述されていることを特徴とする映像の階層的管理方法を記述したプログラムを記録した記録媒体。

【請求項12】 請求項11記載の記録媒体であって、前記映像の階層的管理方法を記述したプログラムが、前記見出し項目が整列された整列済み見出し項目リストをインデクスデータとして、映像データと関連付けて蓄積した蓄積内容を取込み、

複数の当該整列済み見出し項目リストの少なくとも一つを選択させ、当該選択された整列済み見出し項目リスト内の見出し項目に対応する論理単位について、当該論理単位の開始時間から蓄積された映像データを再生する処理過程を有することを特徴とする映像の階層的管理方法を記述したプログラムを記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、映像に関する様々な情報（タイトル、著作権情報、出演者、場所、出来事等）をコンピュータに読み取り可能な形で記述、管理し、効率的に検索・閲覧することを可能にするための映像の階層的管理方法、階層的管理装置並びに階層的管理プログラム記録媒体に関する。

【0002】

【従来の技術】（従来技術1）映像の内容に関する記述

をコンピュータで管理するためのデータモデルとして、成層モデル（stratification model）が提案されている（T. G. Aguiere Smith and G. Davenport, "The stratification system: A design environment for random access video," Workshop on Network and Operating System Support for Digital Audio and Video, Nov. 1992）。

【0003】当該成層モデルによる映像内容記述の例を図7に示す。時間軸上に、層（stratum）と呼ばれる記述を重ね合わせていくことで、映像内容を記述する。層801は、10～40秒の区間に「四季」というタイトルの音楽が挿入されていることを表し、層802は、ナレーションが15～45秒の区間に挿入されていることを、層803は35～42秒の区間に富士山の映像が入っていることを、層804は25～40秒の区間に「日本の四季」という字幕が入っていることをあらわす。このように、重なりや包含を許す層を重ねることで複雑な記述を可能にする点が特徴であった。

【0004】成層モデルで記述したデータを用いると、例えば、「富士山」の写ったシーンをデータベースから検索することが可能となる。さらに、「富士山が写っているシーン」かつ「音楽の流れているシーン」を検索する（35～40秒の区間が該当する）ことも可能となる。時間情報が得られるので、映像がデジタル化されていればその時刻に素早くアクセスすることができる。

【0005】大量の映像を管理する場合、キーワードによる検索だけでなく、カテゴリーに映像を分類し、カテゴリー選択→ビデオタイトル選択→…という具合に段階的に対象を絞っていく階層ブラウジング方式が有効である。従来のシステムではビデオタイトルを選択すると、30分から1時間もある映像を頭から再生するものがほとんどで、すべてを見る前に飽きて見るのをやめることが多かった。

【0006】教育用ビデオ、ドキュメンタリ、研修用ビデオ、ニュースには、本の章節にあたる論理構造が存在する。映像と関連付けて、コンピュータが読み取り可能な形で論理構造が記述されていれば
カテゴリー選択→ビデオタイトル選択→目次を表示し、
章または節を選択→映像再生

といった形で、利用者が見たい場面に効率的にアクセスできるようになる。例えば、「日本の歴史」というタイトルのビデオがあれば、その目次が例えば図6に示す如く表示され、利用者が弥生時代に興味があれば、弥生時代の項を指示選択し、縄文時代の説明を飛ばして弥生時代に関する説明だけを見ることができる。

【0007】従来技術1の成層モデルは、階層構造を記述できないため上のような階層ブラウジングを実現できなかった。

【0008】キーワード検索だけでなく、このような階層的なブラウジングを可能にするには、映像と関連付け

て論理構造（上の例で目次にあたる情報）を記述する仕組みが必要になる。

【0009】（従来技術2）映像の論理構造を記述する方法が特開平5-282379「動画像の管理方法及び管理装置」に開示されている。この方法は論理構造を図8に示すような木データ構造で管理する。まず映像を時間的に重なるの無い区間（C1, C2, C3, …）に分割しそれを葉ノードとする。多くのシステムではショット、すなわちカメラのON/OFF操作をはさまないで連続して撮影された区間に映像を分割しそれを葉ノードとする。従来法は、複数の葉ノードを手で意味的なまとまりに併合して上位ノードS1を作成するという手続きを、ボトムアップに繰り返すことで論理構造を木データ表現する。あるいは、ノードを複数の子ノードに分割するという手続きをトップダウンに繰り返すことで木データ構造を構築する。

【0010】映像の論理構造を表現するのに木データ構造を利用すると、以下の二つの問題が生じる。

（1）各ノードが映像区間に対応するため、映像の時間的線形性に起因する制約を満たすように操作者は注意深く操作を行う必要がある。例えば、図8の例でC1, C3をノードS1の子ノードに、C2をノードS2の子ノードにするといった時間の入れ替えは禁止される。

（2）「ノードを分割し、子ノードを二つ作る」、「子ノードを併合して親ノードを作る」という操作を繰り返す必要があり、操作者にとって必ずしも直感的とは言えない。例えば図8のようなデータ構造を作成してから、C4のノードはS1の子ノードに移す場合を考えると複雑な手続きを経る必要がある。

【0011】つまり、オペレータが論理階層構造を手手で付与する手続きが複雑であるという問題があった。

【0012】（従来技術3）従来技術1の論理構造を表現できないという問題点を解決するために、成層モデルと木データ構造を組み合わせる方法が考えられる。Carrierら（Marco Carrier, Leonardo Ligresti, Gulrukh Ahanger, and Thomas D. C. Little: An Annotation Engine for Supporting Video Database Population. Multimedia Tools and Applications, Vol. 5, pp. 233-258, 1997）の方法は、最終的な出力フォーマットとして、電子的に文書を交換するためのマークアップ（マーク付け）言語であるSGML（Standard Generalized Markup Language）を用いているが基本的には、成層モデルと木データ構造との組み合わせと考えることができる。成層モデルとして内容に関する記述を管理し、木データ構造として論理構造に関する記述を管理する。この方法も、やはり木データ構造に起因する上述の問題点が存在する。

【0013】

【発明が解決しようとする課題】本発明は、上記問題点を解決するためになされたものであり、（1）映像の論

理的な階層構造を記述可能とする、（2）オペレータの直感的な操作により論理構造を付与可能とする、映像の階層的管理方法、階層的管理装置並びに階層的管理プログラム記録媒体を提供することを目的とする。

【0014】

【課題を解決するための手段】請求項1は、コンピュータを用いて映像の内容を階層的に管理する映像の階層的管理方法であって、映像の論理単位についての、当該論理単位に対応するタグと当該論理単位に対応する映像を表示する開始時刻と当該論理単位に対応して存在し得る見出し文字列とを、1つの見出し項目として指定する処理を繰り返し実行して、見出し項目が配列された見出し項目リストを作成し、当該見出し項目リスト内の各見出し項目について、各見出し項目の前記タグにもとづいて、当該見出し項目に割りつけられる見出しレベルを取得し見出し項目リストを開始時刻の小さい順に、開始時刻が同じであれば見出しレベルが小さい順に見出し項目を整列させ、見出しレベルの大小関係を表示することを特徴とする。開始時刻とともに見出しレベルを指定させ、論理構造に対応する順序に並べ替えて、文章の見出しのように見出しレベルの小さいものは目立つように表示することにより、簡便な論理構造の可視化を可能とする。

【0015】請求項2は、見出し項目リストに記録された開始時刻と見出しレベルを参照して、見出し項目をノードとする階層構造に変換し、各見出し項目を階層的に表示することを特徴とする。階層的な構造を構築することで、論理構造を直接的に表現可能とする。

【0016】請求項3は、映像の論理構造の終了時刻を更に見出し項目に指定することで、映像の再生開始時刻と終了時刻を利用者が指定できるようにし、映像の部分的再生を可能とする。

【0017】請求項4は、映像から抜き出した複数の代表画像を見出し文字列と同時に一覧表示することを特徴とし、見出し文字列だけではなく、映像の画像内容を代表画像を用いて表示できる。

【0018】請求項5は、時間軸上に代表画像と一対一に対応付けられた帯をタイムライン表示し、タイムライン表示された帯を選択すると、対応する代表画像が選択されて当該代表画像が選択されたことが可視的に表示されることを特徴とする。タイムライン表示により、映像に関連付けられたタグ情報の時間長、相対的位置関係が視覚的に確認できるようになる。

【0019】請求項6は、代表画像と関連付けてテキストをはじめとする画像、音声、リンク等の情報を管理することで、映像のキーワード検索、音声変換、ハイパーリンク化、要約等を可能にする。

【0020】請求項7は、見出し項目リストが整列された整列済み見出し項目リストをインデクスデータとして、映像データと関連付けて蓄積しておき、見出し項目

リストの少なくとも一つを選択させ、論理単位の開始時間から蓄積された映像データを再生することを特徴とする。見出し項目リストを、映像の目次として表示することにより、利用者が映像の論理的構成を理解できるようにし、それぞれの論理単位の開始点から映像を再生できるようにする。

【0021】請求項8は、見出し項目リストとともに検索用インデクスを蓄積しておき、利用者が検索キーとして単語、文章、画像、等を与え、該検索キーと一致または類似するインデクスを列挙し、該インデクスと時間的に重なりを持つ見出し項目を見出し項目リストから検索し、目次付きの検索結果を返すことを特徴とする。検索用インデクスを蓄積しておくことにより、キーワード検索等を可能にし、検索結果を利用者に返す前に、目次の情報を付与することで検索結果の理解を容易にする。

【0022】

【発明の実施の形態】本発明の実施形態例について図を用いて説明する。

【0023】実施形態例1

●装置構成

図3は本発明の一実施例の装置構成である。本発明はコンピュータプログラムを記録した記録媒体を装着し、当該記録媒体の内容を読み込んで実行するようにしても実現可能である。

【0024】図3において、主制御部305によって、ディスプレイ301、キーボード302、マウス303、スピーカー304、ハードディスク装置306がバス307などを介して制御されている。ハードディスク装置306には映像データ308とそれに対するインデクスデータ309（詳細については後述）とが保存されている。映像データ308は、MPEG等のデジタルビデオのフォーマットにより符号化されており、ファイルとしてハードディスクに記録されている。

【0025】●インデクスデータ309

インデクスデータ309の具体的な構成を図4に示す。インデクスデータ309はヘッダ部401と本体402とから構成される。

【0026】ヘッダ部401には映像データ全体に関する情報が保存される。例えば、映像タイトル、サブタイトル、映像圧縮形式（画像、音声）、映像長、著作権者（会社名、個人名）、利用条件、分類、キーワード、制作会社、ID番号等が保存される。

【0027】本体402は、映像に人手または自動的に付加される情報—“タグ”（識別札）と呼ぶ—を管理する。図4の本体402の各行が一つのタグに対応する。それぞれのタグは、属性として、タグ種別410、開始時刻411、終了時刻412、代表画像413、テキスト414、その他のデータ415などをもっている。

【0028】タグ種別410はタグとして付加された情報の種別を表す。図示実施例で利用するタグ種別は以下

の6種類である。

1) 大見出し

2) 中見出し

3) 小見出し

…これら3種類のタグは映像の論理構造を表すためのものであり、

4) ショット

5) テロップ（編集段階で字幕が挿入された区間）

6) 音声（人の声が入った区間）

…これら3種類のタグは映像内容を記述するためのものである。

【0029】勿論、これ以外にも映像中に現われる人物、出来事、場所、絶対時間等の情報を表すためのタグを追加することが可能である。

【0030】開始時刻411、終了時刻412の列には、映像ファイルの先頭からの時刻がミリ秒の単位で格納されており、開始時刻と終了時刻とによってタグと映像区間とが対応づけられる。

【0031】代表画像413の列には、タグと関連づけられた映像区間を代表する一枚の画像を指定する。図4の例では、時刻を表す数字が格納されており、必要に応じて代表画像を映像データ308から取得することができる。もちろん、時刻を格納するのではなく、代表画像を格納したファイル名を格納することもできる。このような形態はインデクスデータ309を映像データ308と分離して流通させる場合に好適である。また、ここでは各映像区間にそれぞれ一枚の代表画像を登録するようにしているが、複数枚登録できるようにもできる。

【0032】テキスト414の列には関連するテキストデータが格納される。図4の例では、503行ではテキストとしてテロップ（字幕）の文字列が格納されており、508行のものは人の声に対応するタグであり、ナレーションとして読み上げられた文章がテキストとして格納されている。もちろん、その他のタグ、例えば、ショットをあらわすタグに人物、物、場所、キーワードなどを登録することも可能であり、インデクスデータ309の用途に応じて様々な利用形態が考えられる。

【0033】その他415には、代表画像、テキストとして記述しきれないデータを格納する領域である。例えば、508行ではナレーションを読み上げた人の名前を格納している。字幕の表示位置や、カメラワークのタイプ等、様々なものを格納できる。

【0034】インデクスデータ309は、オペレーティングシステムの管理するファイルとして保存することもできるし、データベース管理システムによって管理するようにしてもよい。

【0035】●インデクス自動作成処理

インデクスデータ309を映像処理技術等を用いて自動的に抽出する方法が様々な開発されている（P. Aigrain,

H.-J. Zhang, D. Petkovic: Content-Based Representation and Retrieval of Visual Media: A State-of-the-Art Review. Multimedia Tools and Applications, Vol. 3, pp. 179-202, 1996)。

【0036】上述したタグ種別のうち、「4」ショット、「5」テロップ、「6」音声」は上記文献に記載された方法により自動的に抽出可能である。

【0037】「1」大見出し、「2」中見出し、「3」小見出しの見出しに対応するタグは、映像の意味的な内容を理解しないと作成できないため、自動作成処理は現状困難であり、オペレータが一つ一つ手作業で情報を付加している。なお、自動付与可能なタグについても、その結果が100%正しいとは言えないため、人手で修正する作業が必要になることも多く、そういった場合にも自動付与されたタグを修正するためにユーザインタフェースを用いることができる。

【0038】●タグ付与を効率化するユーザインタフェース

本実施形態における、ユーザインタフェースについて図1と図2とを用いて詳細に説明する。本ユーザインタフェースは、人手で映像にタグを付加するためのものであり、操作に応じて内部的にはインデクスデータ309が更新されるようになっている。

【0039】図1はコンピュータ画面の模式図である。この画面は、タイムライン部120、画像一覧部130、モニタ140からなる。

【0040】まず、再生モニタ140は映像データ308の再生141、停止142、巻き戻し143、早送り144等を可能にする。再生ボタン141を、マウス303で指示することにより、映像再生処理が起動され、ハードディスク装置306から映像データ308が読み出され、動画像が再生モニタ140に順次表示され、音声はスピーカー304から連続的に出力される。

【0041】タイムライン部120はインデクスデータ309として格納されているタグ情報をタイムライン上に表示している。縦軸が時間に対応する。タグを開始時刻、終了時刻で区切られる一つの帯として表示する。長い区間に対応するタグは長い帯として表示する。タグ種別毎に、色分けすると同時に、別々の列（大見出し100、中見出し111、小見出し112、カット113、テロップ114、音声115）に表示することによって、どの種別のタグであるか一目で分かるように、また重ならないように表示している。再生モニタ140で現在再生している時間的位置がわかるようにカーソル121（横線）をタイムライン部120に表示し、再生モニタ140の再生動作と同期してカーソル121を上から下に移動させる。また、時間軸（縦軸）のタイムスケール（目盛りの細かさ）を調整できるようにする。

【0042】このようにタグをタイムライン上に並べて表示することで、（1）それぞれのタグが関連付けられ

た区間の時間長が一目で分かる、（2）タグとタグとの時間的な関係、例えば、カットが一定時間毎に入っていると、テロップと音声の始まりが一致しているといった関係が一目でわかる、（3）カーソル121を表示することで映像、特に音声とタグとの時間的な同期関係を把握できる。

【0043】画像一覧部130には、代表画像をアイコンとして一覧表示して映像内容を視覚的に把握できるようにする。大見出し1Aは代表画像101とテキスト102とから構成され図4に示す大見出しのタグ501に対応している。代表画像101は列413から、テキストは列414から取得し、画面上に表示する。大見出しは他のアイコンよりも目立つように大きな文字で表示し一行を占めるように配置する。代表画像110は図4に示すタグ502、503、504に対応するものである。中見出し1B、小見出し1Cも一行を占めるように配置するが、文字は大見出しよりも小さいフォントを用いて見出しのレベルを判別できるようにする。ショット、テロップ、音声に対応する代表画像110は、文章における単語のように左から右へ、上から下へ順に配置する。

【0044】初期状態では、すべてのタグが画像一覧部130にアイコン表示されているが、関心のあるタグのアイコンだけが画面上に現われるようにすることが、操作性を高めるために重要である。関心の無い情報が画面上に表示されていると、重要な情報を見落とししたり、重要な情報を探したのに時間がかかるからである。図1のインタフェースでは、画面上部にタグ種別に対応する6つのボタン（151～156）を用意している。このボタンはオペレータのマウス操作でON/OFFの二つの状態が切り替わるようになっており、ボタンがONからOFFの状態になると画面上が当該タグ種別のアイコンが画面上から見えなくなるように再描画を行う。具体的には、ボタン151～156は押し下げられた状態

（ON）と通常の状態（OFF）の二つの状態のいずれかを持っており、状態変化が生じると以下の処理を実行し、画面再描画を行う：（i）まず、ONの状態にあるボタンを列挙しそのタグ種別のリスト（L）を得る、（ii）画面上のすべてのアイコンを消去する、（iii）図4の表402を上から順に見て、タグ種別401がリストLに含まれれば、そのタグに対応するアイコンを画面に表示するようにする。

【0045】画像一覧部130に表示されたアイコンをオペレータがマウス等で指示すると、対応するタグの開始時刻の画像を再生モニタ140に表示し、対応するタグが選択状態になる。タイムライン部120と画像一覧部130とを並べて表示し、タイムライン上に表示された帯（一つのタグに対応する）をマウスでクリックした場合も、画像一覧部130内の対応するアイコンが選択状態になる。さらに、マウスのダブルクリック操作によ

り、再生モニタ140で再生が開始されるようにすると操作性が向上する。より具体的には(i)タイムライン上の一本の帯は一つのタグに対応しており、この対応関係を保持するために、帯を表す矩形(右上と左下隅の座標値)とタグ番号(図4の402の行番号)との対応表を用意しておき、(ii)タイムライン上の帯がマウスで選択されると、上記対応表を参照して、タグ番号を取得し、(iii)画像一覧部130でも同様に対応表を用意しており、タグ番号をキーとしてその対応表をひき、対応するアイコンをハイライト表示するなどして選択状態にする。

【0046】このように文章の体裁を真似た形で見出しを目立つように配置することで、映像の論理構造をオペレータが直感的に理解できるようになる。代表画像を一覧表示することでオペレータは映像内容を視覚的にすばやく把握できる。代表画像の指示で、すばやく特定のタグを選択できて、操作効率が上がる。

【0047】また図2に示すように木構造表示することもできる。大見出し1A'を根とし、その下層に中見出し1B'を配置し、その下層に小見出し1C'、最下層にカット、テロップ等に対応するアイコン110'を表示する。ここでは詳細情報を表示せずタグ種別を表す記号(C, T, ...)だけを表示し、画面内に多くの情報を収めて長時間映像の構造が一目で理解できるようにしている。なおアイコン110'に表示されている記号T, C, C, Cはタグの種別を表しており、Tはテロップ、Cはカットのタグであることを表す。画面スペースが足りなければ、図1のタイムライン部120をオペレータの指示にしたがって階層表示部200に切り替えることができるようにできる。このような表示をするには、木データ構造を構築する必要があるが、その具体的手続きについては後述の「階層構造を構築する方法」の所で述べる。

【0048】(タグ削除)画像一覧部130のアイコンのうち一つまたは複数個を選択状態にし、メニューまたはキー操作で「タグ削除」コマンドを起動する。内部的には、選択状態のアイコンに対応するタグを見つけ、インデクスデータ309の表から該当する行を削除する。インデクスデータ309が更新されると、その結果、図1の画面から選択状態になっていたアイコンが消去される。アイコンが消去されたことによって、空きスペースができるが再レイアウト処理が起動され空きスペースに他のアイコンが再配置される。より具体的には(i)選択状態になったアイコンからタグ番号を取得し、図4の表402からそのタグ番号に対応する行に削除されたことを示すマークをつけ、(ii)画像一覧部130、タイムライン部120、階層表示部200を再描画する。

【0049】(タグ追加)画像一覧部130のアイコンを一つ選択するとモニタ上にその時刻の画像が表示される。メニューまたはキー操作により「タグ追加」コマン

ドを起動する。内部的には、インデクスデータ309の表に新規の一行を追加することになる。タグ種別410はオペレータに選択させるか、予め指定しておく。開始時刻411は現在モニタ上に表示された画像(選択状態にあるアイコンの開始時刻に一致する)に設定される。終了時刻412は、タグ種別410によって設定方法が異なるが、タグ種別が「カット」であるタグについては、時間的に次のタグの開始時刻を新規タグの終了時刻として設定する。その他のタグについては、ユーザが指定しない限り映像の終わりを終了時刻として仮に設定する。代表画像413は開始時刻の画像とする。テキスト414は空文字列に初期化する。具体的には、図4の表402の最後に一行を追加し、タグ種別、開始時刻、終了時刻を設定し、画面を再描画する。

【0050】このように、一覧表示されているアイコンをマウス等で指示するだけで、開始時刻の指定がなされるので、再生モニタ上で早送り、巻き戻し操作を繰り返して微調整を行う手間が省け、省力化が可能である。

【0051】(タグ併合)画像一覧部130のアイコンを複数個選択した状態で、メニューまたはキー操作により「タグ併合」コマンドを起動する。内部的には複数のタグを一つにまとめる操作になる。もし、選択されたタグのタグ種別がすべて同じであるかを検査し、一つでも異なる種別のタグがあればオペレータにエラーを通知する。開始時刻411、終了時刻412は次のように設定する。選択された複数のタグの開始時刻、終了時刻を順に調べ、その最小値、最大値を求め、最小値を開始時刻に、最大値を終了時刻とする。代表画像413は、選択されたタグのうち先頭のもの代表画像で代用する。テキスト414は、選択されたタグのテキストをつなげたものにする。具体的には、選択されたタグに対応する図4の表402の行をすべて削除済みとしてマークし、表402の最後に一行を追加して併合後のタグに関する情報を格納する。

【0052】(テキスト入力)画像一覧部130のアイコンを一つ選択した状態で、メニューまたはキー操作により「テキスト入力」コマンドを起動する。内部的には選択されたアイコンに対応するタグのテキスト414を編集対象とする。コマンドが起動されると、図12に示すようなウィンドウが新規に表示され、テキスト編集領域1301にすでに格納されているテキストが表示される。テキスト編集領域1301を通してキーボード302を用いてテキスト414の内容が編集可能になる。変更は即座に画面上に反映される。図12に示すように、テキストとして公開用、非公開用といった情報の種類に分けてテキストを入力できるようにすることも、応用によっては効果がある。ボタン1303、ボタン1304は直前、直後のアイコンに移動するためのアイコンであり、テキストを次々に入力できるようにすることがオペレータの操作効率を上げるために効果的である。また、

ボタン1305は、映像の再生を指示するためのボタンであり、再生モニタ140が前面に現われ、該当する映像区間が再生、あるいは繰り返し再生される。このように、テキスト入力を行いながら、映像を確認できるようにすることが好適である。ボタン1306、ボタン1307はテキスト入力を完了、またはキャンセルする。

【0053】（見出しタグ追加）見出しタグも上述したタグ追加手続きとほぼ同じである。オペレータは、一つまたは複数のアイコンを選択した状態で、メニューまたはキー操作によって「見出し追加」コマンドを起動する。ここで、見出しレベルを指定する「大見出し」、「中見出し」、「小見出し」の三つから一つを選択し（見出しレベルを指定することに相当する）、見出し文字列を入力することで、見出しタグを新規作成できる。また、見出し文字列の修正は上述した「テキスト入力」コマンドを呼び出すことで可能である。見出しレベルを任意の時点で変更できるようにすることも簡単にできる。

【0054】このように、見出し文字列を追加する操作は、ワープロで文章に見出しをつけるのと感覚的に同じでありオペレータにとってなじみやすい。

【0055】●階層構造を構築する方法

インデクスデータ309は、もともと図4のような表形式であって木データ構造にはなっておらず、図2のような形式で表示するには、階層的な木データ構造に変換する必要がある。

【0056】図9はフロー図を示しており、図11の木データ構造の模式図を用いて変換手続きについて説明する。

【0057】（a）ステップ1001： 根だけからなる木データ構造（図11の1201）を作成し、現在着目しているノードを示すポインタcnode をrootに設定する。

【0058】（b）ステップ1002： まずインデクスデータ（図4の本体部分402）を次の順序で整列させる。整列の順序は、開始時刻（411）が小さい順に並べ、開始時刻が同じであれば見出しレベルが小さい順に並べる。見出しレベルは、図10のタグ種別と見出しレベルの対応表を参照することにより取得する。大見出しは見出しレベル1に対応し、図2に示すような木データ構造において、1段目に配置することを意味している（根ノードを0段目とみなす）。この整列処理は、上述したユーザインタフェースにおいてオペレータが新規タグを追加したり、既存のタグを消去したり、開始時刻を変更したりした場合に部分的な整列処理が実行されるようにし、つねにインデクスデータは整列した状態に保っておくことが望ましい。図4の本体402は整列された状態になっており、開始時刻は小さい順に並んでおり、開始時刻が等しい場合（501、502の場合など）には見出しレベルの小さい順に並んでいることが確認でき

る。

【0059】図5は、図4に示されたデータ構造をタイムライン表示したものである。時間軸を横軸とし、縦軸に見出しレベルを表示しており、一つの帯（501、502、…）が一つのタグ、すなわち図4の本体402の一行に対応している。

【0060】タグを以下のような順序で整列させる：原点（時刻0）を通り、時間軸に垂直な直線を考える。その直線を原点から右方向に移動させていき、タグを表す帯とぶつかる順（すなわち、開始時刻の小さい順）に整列させる。ただし、二つ以上の帯が左端でその直線と接する場合には、上から順に（見出しレベルの順）に整列させる。図5の例では、直線は原点で大見出し501とカット502と同時に接点を持つので、501、502の順に整列され、続いてテロップ503とぶつかるので503を並べる。このような手続きを繰り返すことで、すべてのタグが規定の順序に整列される。

【0061】なお整列処理の直前に開始時刻を微調整することで整列処理を簡略化できる。見出しレベルをlevel（level = 1, 2, 3 ; level = ∞の場合はlevel = 4と読み直す）とし、開始時刻 start を次式で微小修正する。

【0062】 $start = start + \varepsilon * level$

ただし ε は微小な正の数である。実施例では、時間の単位としてミリ秒をとっており、開始時刻（これは映像をどのタイムコードから再生するかを表している）を1ミリ秒前後させても人は気づかない。従って ε を1にとることができる。この微調整により、整数をキーにした整列処理に帰着され、処理が簡略化されるとともに高速化できる。

【0063】（c）ステップ1003： すべてのタグについて以下の処理を実行するが、すべてのタグを処理し終わったら終了する。

（d）ステップ1004： 見出しレベルを取得し変数level に格納する。図4の例では、1行目（501）は、タグ種別が、“大見出し”であるので、見出しレベルlevel = 1である。

【0064】（e）ステップ1005： level が∞かどうか（つまり、見出しタグか否か）を調べ、もし見出しタグ以外（yes）であれば、ステップ1006の処理を行い、見出しタグの場合（no）にはステップ1007以降の処理に移る。

（f）ステップ1006： 単純に現ノードcnode の子ノードにタグを追加する。

【0065】（g）ステップ1007： 木データ構造はrootから順にたどり、（level - 1）段目の最後尾にあたるノードをcnode として新たに設定する。

（h）ステップ1008： cnode の子ノードにタグを追加する。

（i）ステップ1009： 今追加した子ノードを現ノ

ードcnodeとして記憶する。

【0066】図4の例と、図11の模式図とを使って動作を説明する。まず、木データ構造の初期状態は図11図示1201のようになっており、root以外にはノードが無い状態である。

【0067】501行目を見ると大見出しのタグ（見出しレベル=1）があるので、ステップ1007ではレベル0のノード（つまりroot）をcnodeとし、ステップ1008では大見出しタグ（501行）をその子ノードとして追加し、ステップ1009ではcnodeをレベル1に追加したノードを指すようにする（図11図示1202）。

【0068】次いで502～504行を見ると、見出し以外のタグが並んでいるので、図10の対応表から見出しレベルは∞で、これら三つのタグは現ノードcnodeの子ノードとして木データ構造に追加され、図11図示1203の状態になる。

【0069】505行を見ると子見出しのタグ（見出しレベル=2）があるので、レベル2のノードとして追加され、現ノードcnodeがそのタグを指すようになり、図11図示1204の状態になる。

【0070】506～508行は、見出し以外のタグであるからcnodeの子ノードとして追加される（図11図示1205）。

【0071】図4には表示されていないが、次に中見出しタグとして、図2に示す如く「米国の四季」というタグが現われると、図11図示1205中の点線で結ばれているノードが追加されることになる。

【0072】なお、以上のような手続きで階層構造を構築すれば、映像の先頭から章節の番号（例えば、大見出しには1, 2, 3, …、中見出しには1.1, 1.2, …）を自動付与することが可能となる。自動付与された章番号は、図1の画像一覧部130に示す見出し文字列102と一緒に表示することができ、章番号を表示することで論理構造を利用者が把握しやすくなる。

【0073】●映像閲覧

図13はシステム構成図を示し、図14のフロー図と図15の画面遷移図とを用いて、映像データベース閲覧の一実施例について述べる。

【0074】図13のシステム構成図は映像データベースの閲覧、検索を可能にするためのものである。ユーザ側のクライアントコンピュータ1404はネットワーク1403を介して、検索サーバ1402、映像サーバ1405に接続されている。検索サーバ1402は、上述した方法で作成したインデクスデータを複数の映像について蓄積した蓄積装置1402と接続されている。映像サーバ1405は、ネットワークで配信可能な映像フォーマットに変換された映像データが複数蓄積された蓄積装置1401に接続されている。もちろん、映像サーバ、検索サーバは複数あってもよく、検索要求の数に応

じて自由に設計できる。

【0075】図14のフロー図と図15の画面遷移図とを用いて、閲覧手続きについて説明する。まず、検索サーバ1402は、蓄積されている映像データのタイトル一覧をクライアントに送信する（ステップ1501）。

【0076】クライアント画面上に表示されたビデオタイトル一覧（1601）の中から、ユーザが一つのビデオタイトルを選択する（ステップ1502）。

【0077】検索サーバ1402は、利用者が選択したビデオタイトルから、上述した手続きにより目次データを、インデクスデータ1402から作成しクライアントに送信する（ステップ1503）。

【0078】クライアント画面に図15図示の画面1602に示すような目次画面が表示される。ユーザは目次画面1602を見て、一つの項目を選択する（ステップ1504）。

【0079】図15図示の画面1602には各目次項目毎に「PLAY」、「詳細」という2種類のボタンが併記されており、「詳細」ボタンが押下された場合には、ステップ1505に移る。

【0080】ステップ1505は、検索サーバに問い合わせ、該当する目次項目の下位にあるショットを検索し、その代表画像をクライアントに送信し、クライアント画面に図15図示の画面1603に示すような画像一覧を表示する（ステップ1505）。

【0081】図15図示の画面1603に表示された画像または見出し文字列をユーザが指示すると、映像サーバ1405に接続し、指示された代表画像、または見出し文字列からインデクスデータ1402に記録されている再生開始時間を取得し、映像データをその位置からクライアントに送出し、クライアント画面（図15図示の画面1604）には受信した映像データが再生される（ステップ1506）。このように、目次を表示することにより、効率的な閲覧が可能になる。

【0082】●映像検索

図16はフロー図を示しており、図17の画面遷移図を用いて映像検索手続きについて述べる。

【0083】まず、利用者がクライアント側でキーワードを入力する（画面1801、ステップ1701）。なおキーワードとしては、単語あるいは文章あるいは画像のいずれか1つまたは複数の組合せたものからなるが、その他にも、所望の検索を行うに必要な音声や、映像固有の条件例えば著作権情報、時間表などを用いることができる。

【0084】検索サーバはキーワードをクライアントから受け取り、キーワードにマッチするタグをインデクスデータから検索する。具体的には、図4の本体402で、テキスト414の列を上から順に調べることになる。なお、複数の映像がある場合にはそれぞれの映像について作成されたすべてのインデクスデータを順に調べ

る（ステップ1702）。

【0085】時間的に重なりを持つ、見出し項目を列挙する。図5の例で説明すると、「富士山」のキーワードで検索すると、図5に示す「富士山」を示すショットがキーワードにマッチするタグとして見つかる。ステップ1703で、そのタグに時間的に重なりを持つ見出し項目を調べると、大見出し「世界の四季」（501）、中見出し「日本の四季」（505）、小見出し「春」の三つが見つかる（ステップ1703）。

【0086】検索結果をクライアントに返し、画面1802に示すような目次付きの検索結果一覧を表示する（ステップ1704）。

【0087】目次付きの一覧を表示することで、どのような文脈で検索結果が用いられているかが分かる。

【0088】以上の手順をコンピュータによって実行するためのプログラムは、コンピュータが読み取り可能な可搬記録媒体メモリ、半導体メモリ、ハードディスクなど適当な記録媒体に格納することができる。

【0089】

【発明の効果】本発明によれば、（1）映像の論理的な階層構造を記述可能になる、（2）オペレータの直感的な操作により論理構造を付与可能になる、などの効果がある。

【0090】請求項1の発明は、開始時刻とともに見出しレベルを指定させ、論理構造に対応する順序に並べ替えて、文章の見出しのように見出しレベルの小さいものは目立つように表示することにより、簡便な論理構造の可視化を可能とする効果がある。

【0091】請求項2の発明は、階層的な構造を構築することで論理構造を直接的に表現可能とする効果がある。

【0092】請求項3の発明は、映像の論理構造の終了時刻を更に見出し項目に指定することで、映像の再生開始時刻と終了時刻を利用者が指定できるようにし映像の部分的再生（終了時刻に達したら自動的に再生を停止する）を可能とし、プレゼンテーションを行う場合などに余計な部分を見せないようにできる効果がある。

【0093】請求項4の発明は、映像から抜き出した複数の代表画像を見出し文字列と同時に一覧表示することの特徴とし、見出し文字列だけではなく、映像の画像内容を代表画像を用いて一覧表示できる効果があり、これによって映像をいちいち再生して見なくても映像の概略が把握できる効果がある。

【0094】請求項5の発明は、映像に関連付けられたタグ情報の時間長、相対的位置関係が視覚的に確認できるようにする効果がある。

【0095】請求項6の発明は、代表画像と関連付けてテキストをはじめとする画像、音声、リンク等の情報を管理することで、映像のキーワード検索、音声変換、ハイパーリンク化、要約等が可能となる効果がある。

【0096】請求項7の発明は、見出し項目リストをインデクスデータとして映像データと関連付けて蓄積しておき、見出し項目リストを映像の目次として表示することにより、利用者が論理単位の開始時刻からの映像再生を直接的に指示でき、効率的な映像閲覧が可能になる効果がある。

【0097】請求項8の発明は、見出し項目リストとともに検索用インデクスを蓄積しておくことで部分映像を検索が可能となり、目次付きの検索結果を返すことで部分映像がどのような文脈で現われたかを利用者が理解しやすくなる効果がある。

【図面の簡単な説明】

【図1】タグ付与を効率化するためのユーザインタフェースの模式図である。

【図2】階層的表示の例である。

【図3】本発明の一実施形態を説明するシステム構成図である。

【図4】インデクスデータのデータ構造を説明するための図である。

【図5】インデクスデータを説明するための図である。

【図6】画面に表示される「目次」の態様を示す図である。

【図7】従来技術（成層モデル）を説明するための図である。

【図8】従来技術（木データ構造）を説明するための図である。

【図9】階層構造を構築する手続きを説明するための図である。

【図10】タグ種別と見出しレベルの対応表である。

【図11】階層構造を構築する手続きを説明するための図である。

【図12】テキスト入力画面である。

【図13】映像検索システムと映像閲覧システムのシステム構成図である。

【図14】閲覧手続きを説明するための図である。

【図15】閲覧時の画面遷移図である。

【図16】検索手続きを説明するための図である。

【図17】検索時の画面遷移図である。

【符号の説明】

100：大見出し

101：代表画像

102：テキスト

111：中見出し

112：小見出し

113：カット

114：テロップ

115：音声

120：タイムライン部

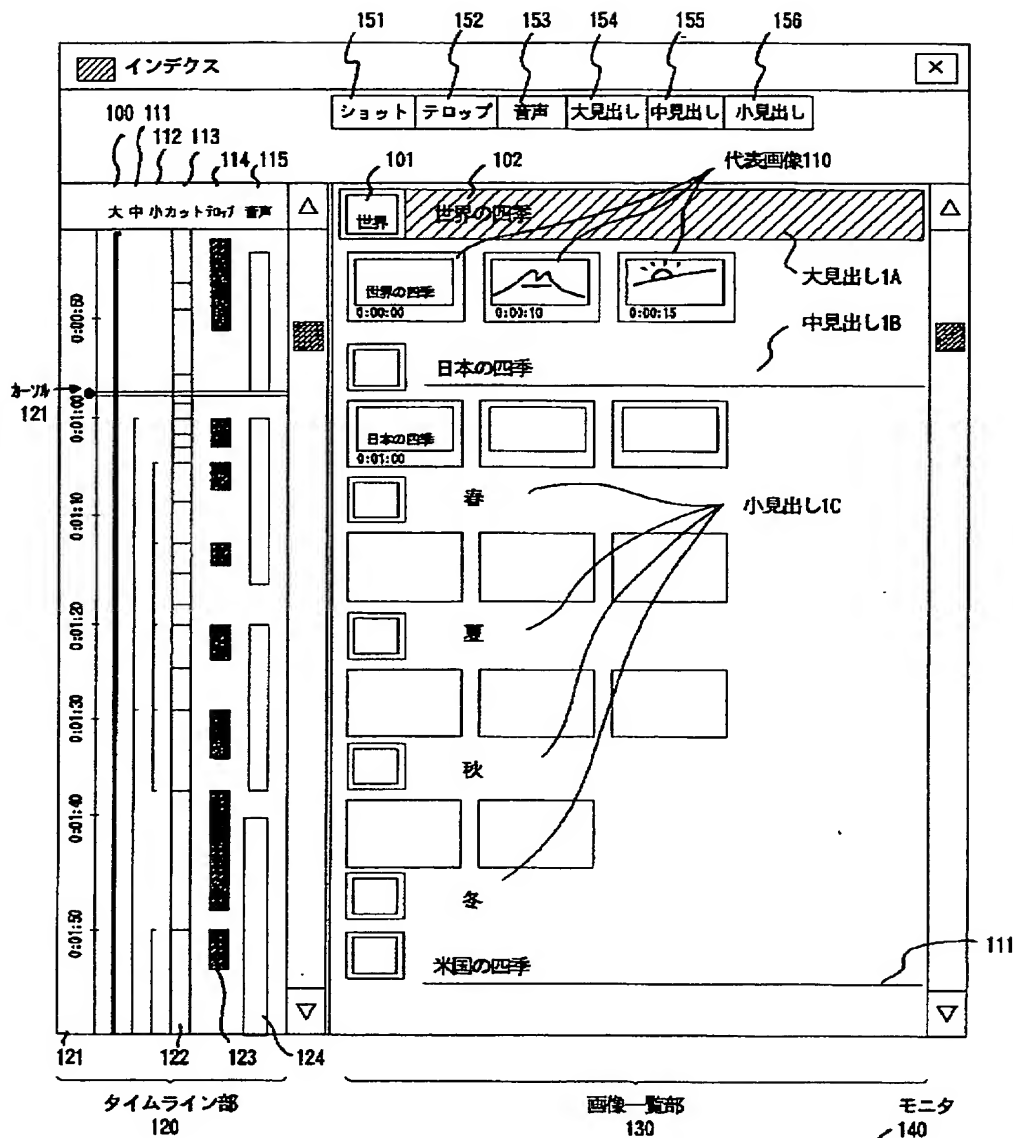
121：カーソル

130：画像一覧部

140 : 再生モニタ
 303 : マウス
 304 : スピーカー

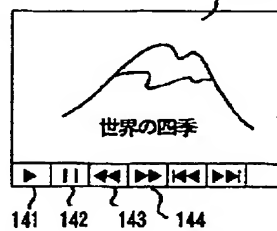
306 : ハードディスク装置
 308 : 映像データ
 309 : インデクスデータ

【図1】

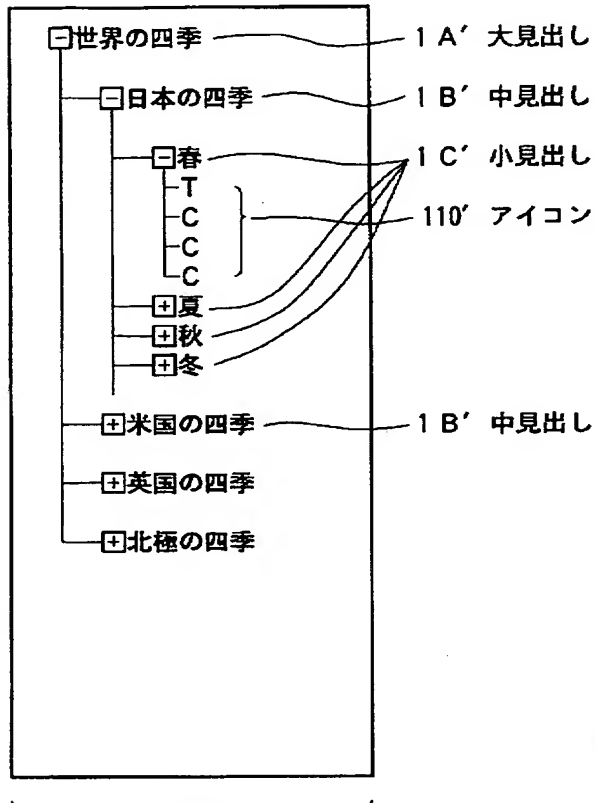


【図10】

タグ種別	見出しレベル
大見出し	1
中見出し	2
小見出し	3
カット	∞
テロップ	∞
音声	∞

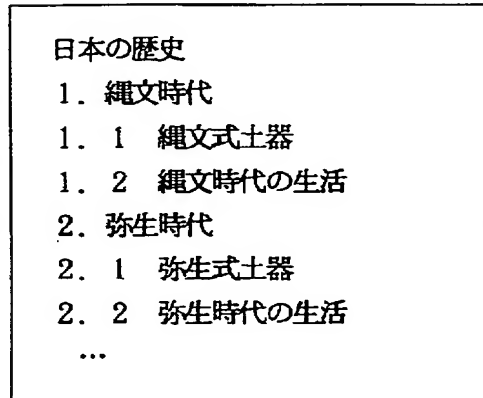


【図2】

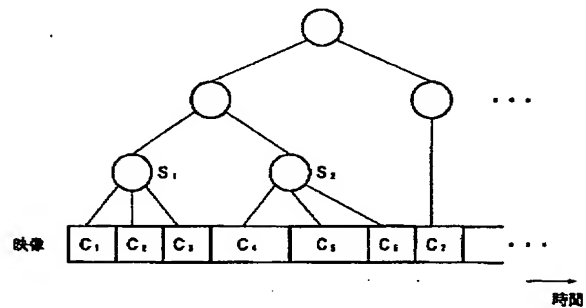


階層表示
200

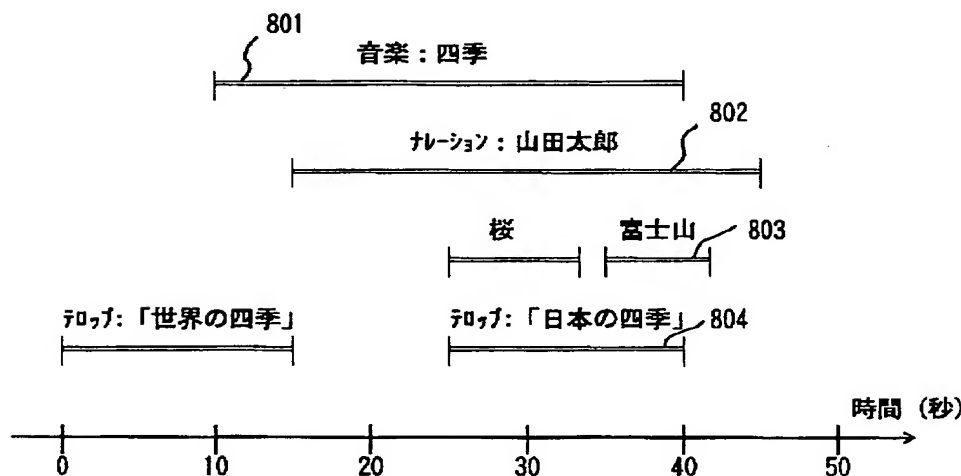
【図6】



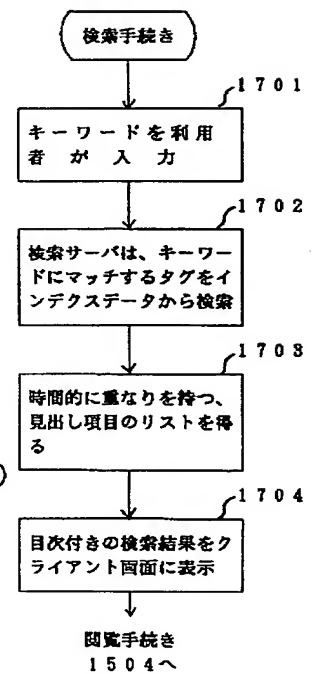
【図8】



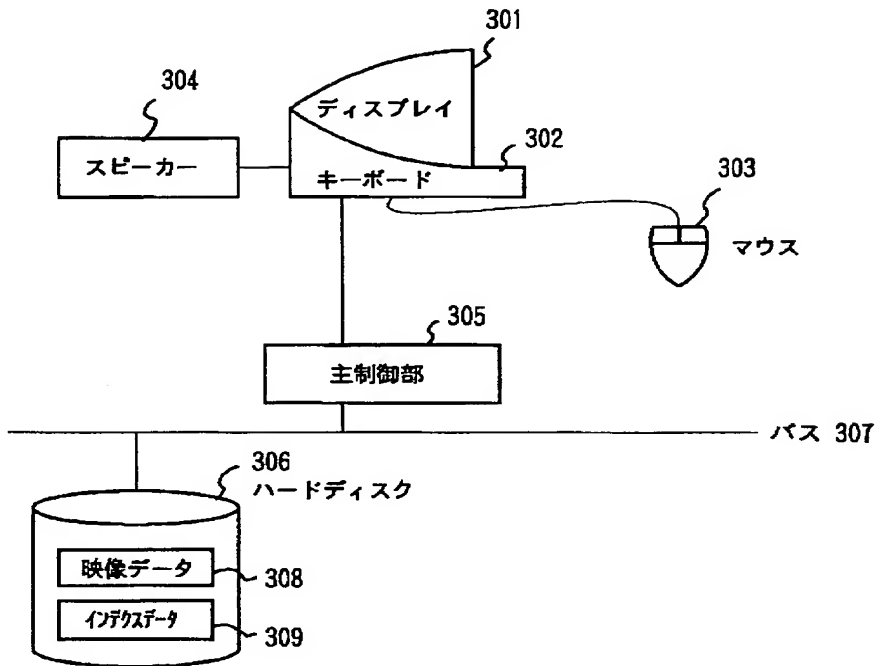
【図7】



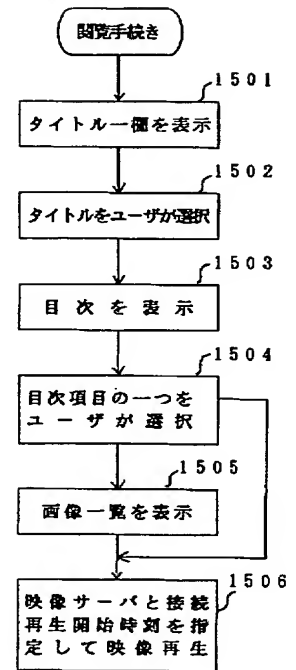
【図16】



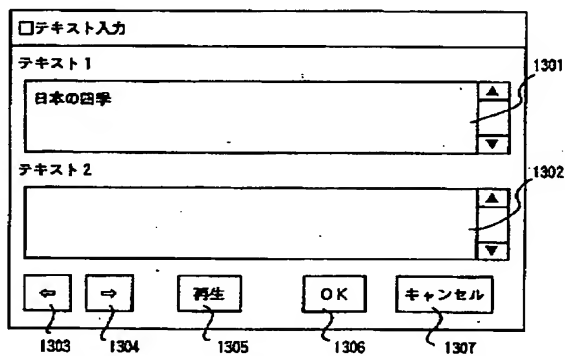
【図 3】



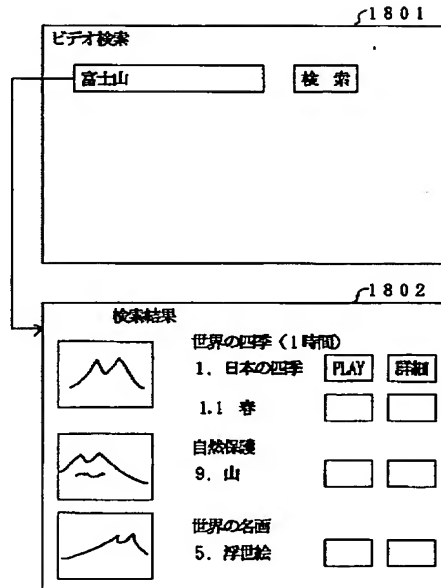
【図 14】



【図 12】



【図 17】

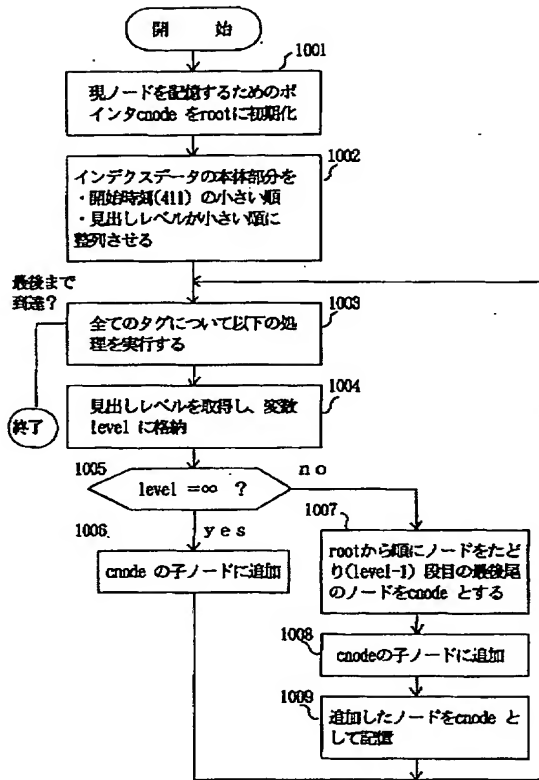


【図4】

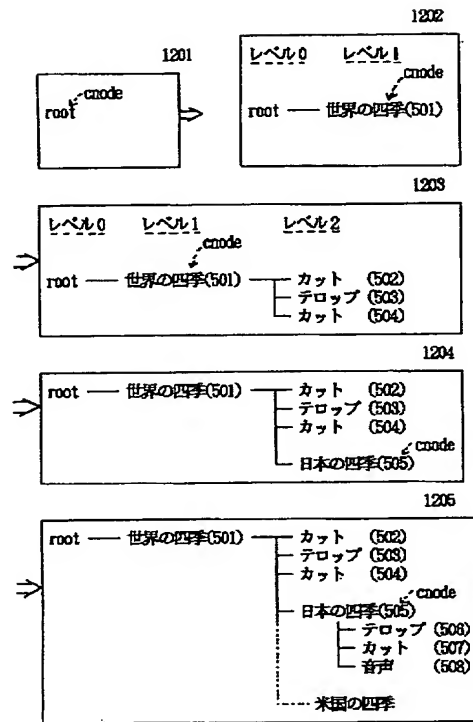
ヘ ッ ダ 部 401	ファイル名	"Seasons. mpg"
	タイトル	世界の四季
	サブタイトル	春夏秋冬の風景
	著作権	〇〇株式会社
	映像長	3, 600, 000
	⋮	

	410 タグ種別	411 開始時刻	412 終了時刻	413 代表画像	414 テキスト	415 その他	
本 体 402	大見出し	0	3, 600, 000	0	世界の四季		← 501
	カット	0	1, 000	0			← 502
	テロップ	500	1, 800	500	世界の四季		← 503
	カット	1, 000	2, 500	1, 000			← 504
	中見出し	2, 500	15, 000	3, 200	日本の四季		← 505
	テロップ	2, 500	3, 500	2, 500			← 506
	カット	2, 500	4, 000	2, 500			← 507
	音 声	2, 500	80, 000	2, 500	このビデオは…	山田太郎	← 508
	⋮						

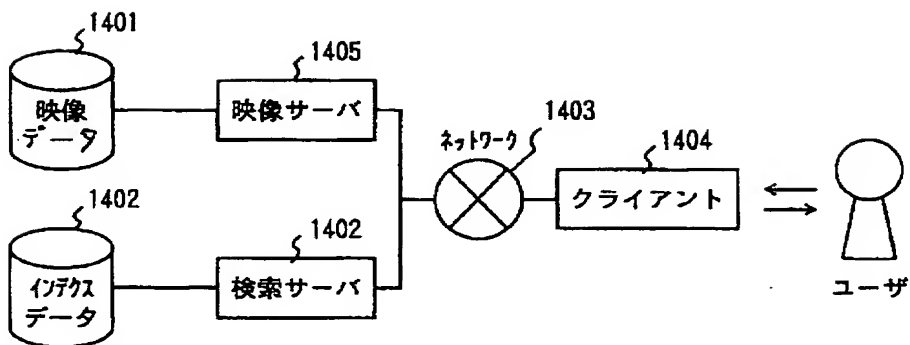
【図9】



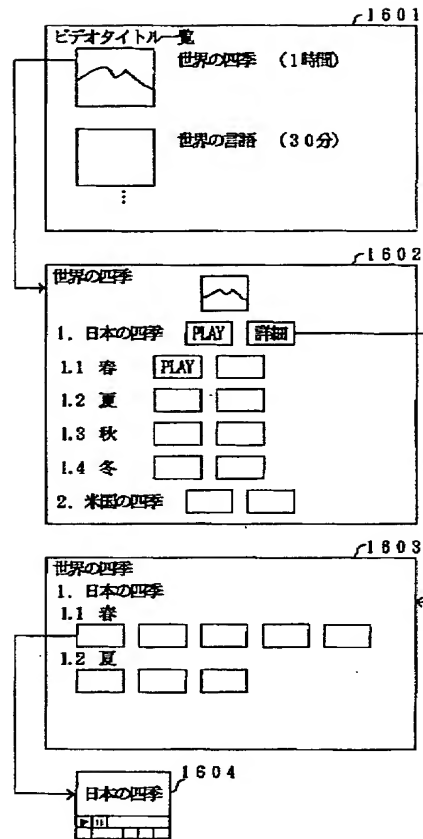
【図11】



【図13】



【図15】



フロントページの続き

(51) Int. Cl. 7	識別記号	F I	テーマコード (参考)
G 0 6 F 12/00	5 0 5	G 0 6 F 12/00	5 0 5
	5 2 0		5 2 0 J
	5 4 7		5 4 7 D
H 0 4 N 5/76		H 0 4 N 5/76	B
5/765		5/781	5 1 0 L
5/781			

Fターム(参考) 5B075 ND12 ND35 NK10 PP03 PP13
 PQ02 PQ13 PQ46
 5B082 AA13 BA03 EA07 GC04
 5C052 AA03 AB04 AB05 AC08 DD04
 EE03